The Unmasking of English Dictionaries

When we look up a word in a dictionary, we want to know not just its meaning but also its function and the circumstances under which it should be used in preference to words of similar meaning.

Standard dictionaries do not address such matters, treating each word in isolation. R. M. W. Dixon puts forward a new approach to lexicography that involves grouping words into 'semantic sets', to describe what can and cannot be said and providing explanations for this. He provides a critical survey of the evolution of English lexicography from the earliest times, showing how Samuel Johnson's classic treatment has been amended in only minor ways. Written in an easy and accessible style, the book focuses on the rampant plagiarism between lexicographers, on ways of comparing meanings of words, and on the need to link lexicon with grammar. Dixon tells an engrossing story that puts forward a vision for the future.

R. M. W. DIXON is Professor and Deputy Director of the Language and Culture Research Centre at James Cook University, North Queensland and a Fellow of the British Academy. His books include *Australian Aboriginal Words in English: Their Origin and Meaning* (2nd edition 1995); *A Semantic Approach to English Grammar* (2nd edition 2005) and *Making New Words: Morphological Derivation in English* (2014). He is the author of the classic three-volume text *Basic Linguistic Theory* (2010–2012) and has published grammars of languages from Amazonia and Fiji, and of several of the original languages of Australia.

Books by R. M. W. Dixon

BOOKS ON LINGUISTICS
Linguistic Science and Logic
What *is* Language? A New Approach to Linguistic Description
The Dyirbal Language of North Queensland
A Grammar of Yidiñ
The Languages of Australia
Where Have All the Adjectives Gone? And Other Essays in Semantics
and Syntax
Searching for Aboriginal Languages: Memoirs of a Field Worker
A Grammar of Boumaa Fijian
A New Approach to English Grammar, on Semantic Principles
Words of Our Country: Stories, Place Names and Vocabulary in Yidiny
Ergativity
The Rise and Fall of Languages
Australian Languages: Their Nature and Development
The Jarawara Language of Southern Amazonia
A Semantic Approach to English Grammar
Basic Linguistic Theory, Vol. 1, Methodology
Basic Linguistic Theory, Vol. 2, Grammatical Topics
Basic Linguistic Theory, Vol. 3, Further Grammatical Topics
I Am a Linguist
Making New Words: Morphological Derivation in English
Edible Gender, Mother-in-law Style and Other Grammatical
Wonders: Studies in Dyirbal, Yidiñ and Warrgamay
Are Some Languages Better than Others?
"We Used to Eat People": Revelations of a Fiji Islands
Traditional Village

with Alexandra Y. Aikhenvald
Language at Large: Essays on Syntax and Semantics

with Grace Koch
Dyirbal Song Poetry: The Oral Literature of an Australian
Rainforest People

with Bruce Moore, W. S. Ramson and Mandy Thomas
Australian Aboriginal Words in English: Their Origin and Meaning

BOOKS ON MUSIC
with John Godrich
Recording the Blues

with John Godrich and Howard Rye
Blues and Gospel Records, 1890–1943

NOVELS *(under the name Hosanna Brown)*
I Spy, You Die
Death upon a Spear

EDITOR OF BOOKS ON LINGUISTICS
Grammatical Categories in Australian Languages
Studies in Ergativity

with Barry J. Blake
Handbook of Australian Languages, Vols 1–5

with Martin Duwell
The Honey Ant Men's Love Song, and Other Aboriginal Song Poems
Little Eva at Moonlight Creek: Further Aboriginal Song Poems

with Alexandra Y. Aikhenvald
The Amazonian Languages
Changing Valency: Case Studies in Transitivity
Areal Diffusion and Genetic Inheritance: Problems in
Comparative Linguistics
Word: A Cross-linguistic Typology
Studies in Evidentiality
Adjective Classes: A Cross-linguistic Typology
Serial Verb Constructions: A Cross-linguistic Typology
Complementation: A Cross-linguistic Typology
Grammars in Contact: A Cross-linguistic Typology
The Semantics of Clause-linking: A Cross-linguistic Typology
Possession and Ownership: A Cross-linguistic Typology
The Grammar of Knowledge: A Cross-linguistic Typology
The Cambridge Handbook of Linguistic Typology
Commands: A Cross-linguistic Typology

with Alexandra Y. Aikhenvald and Masayuki Onishi
Non-canonical Marking of Subjects and Objects

The Unmasking of English Dictionaries

R. M. W. Dixon

Language and Culture Research Centre
James Cook University Australia

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

Contents

Abbreviations and Conventions

-	affix boundary
=	clitic boundary (see pages 12–14)
/… /	encloses a phonological form, made up of a string of phonemes (see page 10)
*	reconstructed form in a proto-language
♦	unacceptable sentence or phrase
DAE	Dictionary of American English
ME	Middle English
NP	noun phrase
OE	Old English
OED	Oxford English Dictionary
PIE	Proto-Indo-European (see pages 147–8)
RP	Received Pronunciation (see page 10)

Abbreviations for syntactic functions (see pages 179–80)

A	transitive subject function
O	transitive object function
S	intransitive subject function

At first mention, dictionaries are generally accorded their full title, in italics; for example, *Webster's Ninth New Collegiate Dictionary*. Later mentions generally use a shorter form, with no italics: for example, Ninth Collegiate.

Prologue: The Work in Advance

A dictionary should tell you when to use one word rather than another. A word must not be regarded as an isolated item. It is a node in the structural framework of the language.

Towards the end of the first millennium, Anglo-Saxon scholars working with Latin texts would write the Old English equivalent below a Latin word. These 'glosses' were gathered together, and put in order. For the next 500 years they served as the first dictionaries, Latin to English. For each Latin word, an equivalent in English was provided; this was its 'definition'. In 1587 Thomas Thomas gave, for the Latin word *lassitūdo*, the English meaning *wearines* (this was one of the spellings used at the time).

The first monolingual English dictionaries – commencing in 1596 – dealt just with 'hard words' (those of foreign origin) and explained them in terms of Germanic forms. The second such dictionary, by Robert Cawdrey in 1604, included:

> **lassitude**, wearines

The Latin head word **lassitūdo**, in the bilingual dictionary, had been replaced by the cognate English word, **lassitude**, in Cawdrey's monolingual dictionary. The Romance word *lassitude* – borrowed into English from French in the middle of the sixteenth century – was 'defined' in terms of adjective *weary* and suffix *-nes(s)*, which derives an abstract noun (both are of Anglo-Saxon origin).

A bilingual dictionary gives translation equivalents between two languages. A Latin sentence including the word *lassitūdo* could be translated by an English sentence with *wearines*. A monolingual dictionary is a quite different matter. It aims to elucidate the meaning of a word (the **head word**, shown in bold type here) in terms of other words of the same language.

We may get a single word following the head word, as with '**lassitude**, wearines'. Does this imply that the words *lassitude* and *weariness* have the same meaning, that each could be substituted for the other? Why should we need two words if they mean the same?

In fact, it is *never* the case that two words have exactly the same meaning. There is always a difference – even if sometimes only slight – concerning what

ix

they refer to, the circumstances in which they are likely to be used, and the pragmatic effects of their use. The function of a monolingual dictionary should be to discuss and compare words with similar meanings, explaining when to use one and not the other.

Lassitude is a fascinating word, linking physical and mental conditions – bodily fatigue on the one hand, and lack of interest or enthusiasm on the other. It has partly overlapping meanings with *tiredness* and *weariness*, and also with *langour* and *lethargy*. In order to know when to use *lassitude*, rather than a word of related sense, a dictionary user should be able to consult an account of the similarities and differences of meaning between *lassitude* and its congeners, their grammatical possibilities, and the interpersonal consequences of selecting one word over another.

By and large, the dictionaries we have today fail to provide this information. In the present book, I will show that this is due to a misconceived methodology. Developing and refining the principles of 'glossing', from a thousand and more years ago, has led to the idea that the meaning of a word – considered as a discrete figment – can be adequately rendered by a 'definition'. In some circumstances, it is suggested that the definition could be substituted for the head word in a sentence, with no significant shift in meaning.

Treating each word as a self-contained entity is the FIRST of three radical faults afflicting present-day lexicography. Each word must be placed within its semantic milieu, the set of words with related meanings. And each such set must be located within the interlocking framework of the entire vocabulary. Words must also be linked to grammar; different nuances of meaning may imply different structural possibilities (and vice versa).

For example, *want* and *wish* have compatible meanings; one can say either *I wish to apply for that position* or *I want to apply for that position*. But, alongside *I wish that I had applied for that position*, it is not permissible to say ♦*I want that I had applied for that position*. The grammatical possibilities correlate with meaning differences, and this must be brought out in a dictionary. (*Want*, *wish (for)*, and *desire* are discussed in chapter 12.)

Finish and *cease* both relate to something which was – but no longer is – happening. One can say either *That bank ceased trading last June* or *That bank ceased to trade last June*. However, with *finish* only the first alternative is possible – *That bank finished trading last June* but not ♦*That bank finished to trade last June*. This grammatical difference reflects a fundamental contrast in meaning for the two verbs. (Chapter 3 has a fullish discussion of *cease*, *finish*, and their congeners.)

Lexical meaning and grammatical properties are intertwined; each helping to characterise the other. Grammar cannot be properly studied without close attention to the semantic proclivities of grammatical slots, and the kinds of lexical words which may fill them. Contrariwise, the contrastive meanings of

words (from a certain semantic set) determine their grammatical functions. Each grammatical construction carries a meaning, and each lexical word has a meaning; these will combine together to show what can be said, and why.

We began by stating that a dictionary should assist the reader in deciding when to use one word rather than another. This aim cannot be achieved through just a list of lexical words, generally in alphabetical order, with a short 'definition' provided for each (and an indication of whether the word is a noun, verb, adjective, or whatever). Lexicon and grammar are interlocking facets of language. In order to learn how to speak and understand a language, equal consideration must be paid to these two components, and to their semantic linkages. Not paying anything like adequate attention to grammar constitutes the SECOND major failing of present-day lexicography.

We can loop back to the early monolingual dictionaries. In the seventeenth century, the head words were forms which had been borrowed from French and Latin during the previous couple of centuries. They were defined in terms of a quite different set of words, those inherited from Old English (plus a few words taken over in the Middle Ages from other Germanic tongues). At that time, the purpose of a dictionary was to explain 'hard words'.

At the beginning of the eighteenth century, the concept expanded. A dictionary should now cover all words in the language, even common adjectives such as *little* and *small*. That is, Germanic words were now included, and they were defined in terms of … well, in terms of other Germanic words. Nathan Bailey's dictionary from 1730 was the best of its era. But some definitions for the most frequent words were not really very helpful. For example:

> **little,** small
> **small,** little in Size, or in Number

Little and *small* may be used interchangeably in some contexts, but by no means in all. In Bailey's time, as today, *little sister* meant younger sister (irrespective of size). One may say, *Only a little snow fell last night*. *Small* is not permissible here. On the other side of the coin, only *small* may be used in *The dresses were all of a small size* and *Only a small number of people were killed in the avalanche*. (Chapter 6 deals with *little* and *small* plus, of course, *big* and *large*.)

Beside being only partial, these two entries of Bailey's are uninformative. If the reader was already familiar with one of the words *little* and *small*, they would get a rough idea of the meaning of the other. If they knew neither, they would simply be bewildered.

Things haven't improved all that much over the centuries. A large modern dictionary lists the following lead senses:

> **choose**, select out of a greater number
> **select**, choose, esp. as the best or most suitable

Select, taken from Latin in the latter part of the sixteenth century, and the Anglo-Saxon word *choose* are here defined in terms of each other.

What the reader needs – to be able to decide what to use when – is explicit comparison of the meanings and functions of the words. *Choose* has a wide general meaning, whereas *select* is a hyponym of it. That is, the meaning of *select* is included within that of *choose*. In virtually every instance, *select* can be replaced by *choose*; for example, *They selected/chose the films to be shown at the festival*. However, there are many uses of *choose* where *select* would not be possible. For instance, when followed by *between*, as in: *She couldn't choose between the diamond ring and the ruby earrings*. And only *choose* can take a TO complement clause, as in: *She chose to close the meeting early*.

Suppose that a foreign learner wanted to know how to use the verb *ask*, and consulted a medium-sized dictionary from the mid-twentieth century. The trail of enquiry goes:

> **ask**, call for an answer to
> **answer**, reply to
> **reply**, make answer, respond, in word or action (to)
> **respond**, make answer (especially of congregation making set answers to priest, etc.); perform answering or corresponding action

The entry for *ask* refers the learner to *answer*. That leads to *reply*, which bounces back to *answer*; but it also mentions *respond*. However, the *respond* entry scarcely helps, simply circling back to *answer*.

If every word can be defined in terms of every other word, where to start and where to stop? One solution would be to recognise a small set of 'basic words' which would not themselves be defined but could be used in the definitions of non-basic words. The definitions would of course have to be succinct, and not umpteen lines long. In the mid-twentieth century there were a handful of dictionaries which adopted a 'defining vocabulary' of 1,500 to 2,000 words; all other words were defined in terms of these. However, words on the defining vocabulary were also accorded definitions and these showed the same old circularity (see pages 204–10).

We mentioned the recurrent suggestion that, in many circumstances, a definition should be substitutable in a sentence for its head word, with no significant shift in meaning. This is sometimes feasible. For example:

> **widow**, woman who has lost her husband by death and has not married again

In place of *Maria is a widow*, one could say, *Maria is a woman who has lost her husband by death and has not married again*. However, this doesn't extend

very far. When our foreign learner hears *Little Jimmy has a new dog* and wishes to know what this is, they look up the dictionary:

> **dog**, any four-legged flesh-eating animal of the genus *Canis*, of many breeds domesticated and wild, kept as pets or for work or sport

The entry is encyclopaedic and instructive, but it can not be substituted into *Little Jimmy has a new –*. As another instance, the definition *call for an answer to* is not substitutable for head word **ask** in, for example, *Tom asked Bill for a cigarette* or *Robin asked Hilary to come to the party*. The same applies for most of the definitions quoted in this volume. I am not suggesting that a definition *should* be substitutable for its head word, but am simply pointing out that this is something which has often been suggested, and that it only occasionally succeeds.

Whether a definition is substitutable or not is a minor matter, something of a distraction. The main point at issue is that, in all modern-day dictionaries, definitions are haphazard, unsystematic and unprincipled. Anything is defined in terms of anything else, and – as with the *ask* and *answer* example – it often leads round in a circle, taking you back to where you started, and not much wiser for the journey.

Relying on definitions is the THIRD shortcoming of present-day lexicography. This ties in with the first one, dealing with each word as an autonomous entity. The way a word is used can only adequately be explained – and understood – by placing it within its semantic and grammatical homeland. This involves matching it against words with similar and opposite meanings, and examining the contexts in which each of these may be used, with well-chosen examples illustrating semantic nuances and subtleties. There will be links from one semantic set to others, and from one grammatical configuration to a related one.

A few years ago, a leading lexicographer explained his plan of attack for a new dictionary – 'Letters A and B will be my responsibility, Ivan will do C, Vanessa D and E, and then I'll chip in again with F, G and H.' A better plan would have been for one person to deal with verbs of motion, another with verbs of speaking, a third with value adjectives, and so on. A set of similar words (for example, *lovely*, *beautiful*, *pretty*, *handsome*, and the like) should be the province of a single lexicographer, who would examine the contrasts between them. As a final step, the entries could be collated into a single alphabetical listing, to make a conventional dictionary.

Constraints on dictionary-making include time, expertise, money (to pay for the time and expertise) and size. As a book, a dictionary has to be of a size that can be afforded and easily used. Space is always at a premium. Each entry is a compromise between informativeness and succinctness.

However, the world has moved on. Computers can accomplish complex tasks in no time at all. During the last few decades, they have been of immense

help to lexicographers in compiling and searching corpuses, laying down a foundation on which definitions are based.

With computers, space limitations dissolve. Things need no longer be presented as a list, in a single dimension. The semantic and grammatical superstructure which characterises a language can be encoded as a multi-dimensional matrix.

Open the 'dictionary file' on your computer, and enter *little*. You will be directed to the semantic set centred on *big* and *little*, *large* and *small* (see chapter 6). There will be sidelinks to sets dealing with *short*, *long*, and *tall*; *deep* and *shallow*; *wide* and *narrow*; and more besides.

There will be characterisation of the words in a semantic set within an overall conceptual template, comparing and contrasting them in terms of structural frames and pragmatic implications. No 'definitions', in the traditional sense; no round-in-a-circle peregrinations.

The plan of this volume is as follows. First, there are chapters on 'How the language is made up' and 'What a dictionary needs to do'. There follows a narrative on the evolution of English dictionaries, step by step – from the early glossing, through dictionaries of 'hard words' to those which dealt with all words (including, inappropriately, grammatical elements such as *the* and *this*). The historical account is interwoven with discussion of types of semantic organisation; with the tradition of copying or part-copying from predecessors (often held to be the only suitable methodology), and with the role of grammar.

Dictionaries came of age with Samuel Johnson, distinguishing different senses of words and including illustrative examples 'from the best writers'. Johnson's methodology underpinned the massive and magisterial *Oxford English Dictionary*, which commenced at the end of the nineteenth century. But then progress halted. During the last century, dictionary-making has been virtually untouched by the emerging science of linguistics. Advances in the understanding of semantics and grammar stand apart, not considered relevant for the grand old tradition of lexicography (treated almost like a branch of history).

Interspersed within the general story are preliminary accounts of four semantic sets – that including *finish*, *cease*, and *stop* in chapter 3*; big* and *little*, *large* and *small* in chapter 6; *fast*, *quick*, *rapid*, *swift*, *slow*, and *speed* in chapter 9; and *want*, *wish (for)*, and *desire* in chapter 12.

The final chapter recapitulates proposals made throughout the book, setting out the blueprint for an innovative – and, indeed, revolutionary – approach to dictionary-making.